REMARKS

Claims 12-18 and 21-30 are pending. Applicants have amended claims 12-18, 22, 25 and 27-29. Claim 24 has been cancelled. No new matter is believed to have been added. Applicants have amended the specification including Figure 1. Figure 1 was amended to correct a typographical error in which an arrow was misplaced. A clean copy of the amended figure is attached herewith. Additionally, the term "alpha," which appears on pages 7 and 8 of the specification, has been removed. A description has been added that summarizes incorporated subject matter described in Patent '247. The following comments are directed to the Examiner's objections in the office action dated December 4, 2003.

Applicants believe it is helpful to summarize the claimed invention before addressing the objections.

The claimed invention presents an important improvement in the field of intein cleavage for protein ligation. The claimed improvement in intein cleavage depends on the novel use of MESNA in place of other thiol reagants such as DTT. MESNA was fortuitously and unexpectedly discovered after significant experimentation with a range of thiol reagents and was not predicted by the applicants or others before the present claimed invention.

The inventiveness of the present claimed invention resides in the use of MESNA as a thiol reagent in intein mediated cleavage for protein ligation and not in the discovery of intein mediated cleavage itself. However, many of the Examiner's comments in the office actions query the mechanisms of protein cleavage. The principles of intein mediated cleavage with respect to

native and modified inteins was discussed in detail and appeared to be understood by Examiners Achatamurthy and Moore at the interview at the U.S. Patent and Trademark Office in August 2003.

The mechanism of cleavage and the use of thiol reagents has been described in previously published patents and references (for example, Chong et al. Gene 192 (1997) reference attached, US 5,834,247 and Telenti et al.). Cleavage reactions rather then splicing occurs under conditions of controllable protein splicing. Controllable splicing occurs when a native intein is in a foreign context, or a mutation prevents splicing absent a particular condition or when an amino acid in the vicinity of the splice junction is chemically modified to result in splicing inhibition subject to a change in a particular condition. If a thiol reagent is present in the reaction mixture when a specified change in condition occurs to induce splicing, cleavage occurs instead of splicing. Splicing is trumped by the presence of a thiol reagent, which induces cleavage of the intein instead of splicing.

Failure on the part of the applicant to respond to any statement by the Examiner concerning his interpretation of the specification is not intended to signify agreement with the interpretation by the Examiner. This latter statement, while not explicitly made previously, is intended to apply to every office action received in this prosecution file.

<u>Rejection under 35 U.S.C.§ 112 first paragraph</u>

(1) The terms " naturally occurring intein", "intein", "intein derivative" and intein mutant" are rejected as not enabled in the methods in claims 12,

17, 25 and 28. The Examiner refers to office action paper 10 , page 5, lines 6-10 and 19-21 for the recitation of reasons for the rejection. However, in the final office action dated 5/23/03, the Examiner appeared to accept applicants' explanation of enablement. The reason is unclear to Applicants as to why the Examiner has now raised the same objection again.

The original objection stated

> Claims 12-16 are rejected for lack of enablement because claim 12 embraces in its clause (a) any and all precursor proteins having a native intein where the specification teaches instead that an intein must be specifically modified by mutating its amino acid sequence or otherwise modifying its structure eg truncating or dividing it or both so that it cannot excise itself with a concomitant protein splicing reaction.
>
> ..where the mutant or derivative thereof is incapable of promoting protein splicing with the amino proximal region of the "precursor protein

In the response dated 12/23/02, applicants stated that support for the use of "naturally occurring inteins" in the method claims could be found for example on pages 6-7 and page 8 of the disclosure.

> The disclosure states on page 8 of the application
>
> " In a preferred embodiment, the intein may be any CIVPS such as SceVMA, Mxe GyrA or derivatives or mutants thereof (referring to Comb et al Patent 5,496,714 and 5,834,247)

Patent 5,496,714 defines a CIVPS as follows:

> CIVPSs of the present invention include any intervening protein sequence in which the splicing reaction can be controlled, either by inherent properties of the native IVPS, such as an increase in

> temperature, or by modifications made to an IVPS that allow the splicing reaction to be controlled. (col 8, lines 59-63)

Patent 5,496,714 explains how a native intein can be controllable when placed in a foreign context or when modified. Patent '247 describes how splicing or cleavage may be initiated under conditions that include temperature, change in pH, exposure to light, phosphorylation, glycosylation and chemical reagents (cols 8 and 12). For clarification, the term "naturally occurring" has been replaced by the word "native". Applicants assert that these adjectives are intended to and actually do mean the same but the use of "native" ensures consistency with the term in the definition used in the '247 patent. Applicant is unaware of any different interpretation of the term CIVPS.

The term " derivative" is further exemplified in Patent '247 cols 8 and 12. A description of "derivative" from the incorporated '247 reference has now been introduced in the specification as specified above following Examiner Nashed's request.

(2) Claims 12, 17, 25 and 28 are found objectionable for reasons laid out in office action paper number 10, page 4, lines 10-12 in which the Examiner developed a theory as to what the specification enabled. In the final office action dated 5/23/03, the Examiner appeared to accept the applicants' arguments that the claims are enabled. This response is presented in (b) on page 7 of the response dated 12/23/02. The reason is unclear to Applicants as to why the Examiner has now raised the same objection again.

The Examiner has rejected the claims again asserting that the specification states a requirement for modifying a naturally occurring intein by at least replacing the C-terminal asparagine. This is incorrect. Applicants used the asparagine modification of the Mxe GyrA intein in Example 1. There is no assertion in the specification that a modified asparagine is required for all inteins. Indeed, applicants explained in the interview in August 2003 to the apparent satisfaction of supervisory Examiner Achutermurthy and Examiner Moore, how the Telenti reference shows that cleavage of a native or modified intein can be achieved without replacement of a C-terminal asparagine.

The Examiner asserts that the specification does not enable "a component other than at least an intein that is <u>incapable</u> of promoting protein splicing at its amino terminal fusion with the selected protein". The Examiner is mistaken. No such requirement is presented in the specification nor is it appropriate. Indeed, applicants explained in the interview in August 2003 to the apparent satisfaction of supervisory Examiner Achutermurthy and Examiner Moore, how the Telenti reference showed that splicing competent inteins can be induced to undergo N-terminal intein cleavage in the presence of DTT.

The Chong reference is provided here to further instruct the Examiner on the mechanism of cleavage in the presence of a thiol reagent.

(3) The subject matter of Claim 24 has now been integrated into claim 22 following the suggestion of the Examiner.

(4) The Examiner has previously rejected the claims to formation of a C-terminal thioester using the thiol reagent DTT in light of Patent '247 (Figure 28 of the application). In contrast to this prior art reference, independent method claims 12, 22, 25 and 28 specifically requires MESNA. Enablement is fulfilled by the teaching in the specification for substituting MESNA for prior art thiol reagents in intein mediated cleavage reactions (page 8 of the application) to produce higher yields of cleavage product suitable for protein ligation.

(5) The Examiner has queried the following:

(a) The relative size of the inactive protein and the chemically synthesized peptide or protein,

Applicants have described in Example 1 on page 14 of the specification how a protein or peptide can be synthesized using standard Fmoc^{tm} chemistry. The claims are not intended to be limited by an absolute size for a chemically synthesized protein or peptide but rather is limited by the capability of synthetic chemistry.

(b) Restoration of protein activity and how to assess the same.

Claim 28 directed to restoring protein activity is enabled in the application by the description in Example VI which describes how a protein exemplified by the restriction enzyme protein-HpaI, can be reconstituted by ligation to form an active protein. It is possible that others (including the Examiner) may make assumptions about protein folding but this is not necessary for the claimed invention which experimentally demonstrates

restoration of protein activity subsequent to ligation of two parts of a protein.

The activity of the ligated protein was assessed in the examples of the above application by using a standard activity assay for the protein which demonstrated that activity was restored. It is generally accepted by those of ordinary skill in the art that the restoration of activity of a protein can be assessed in an activity assay.

Applicants assert that the specification supports the full scope of claim 28 because a person of ordinary skill in the art would be able to reproduce the claimed invention without undue experimentation.

In summary, the applicants respectfully assert that the claims are enabled by the present specification in view of the state of the art at the priority date in 1998 and that the rejection should be reversed.

Indefiniteness

Applicants thank the Examiner for suggestions on how to amend the claims for purposes of clarification. In large part, applicants have adopted the Examiner's suggestions to remove any basis of indefiniteness. None of the amendments are believed to alter the scope of the claims. Applicants respectfully request reversal of the rejection.

Double patenting

Claims 22 and 27-30 are provisionally rejected under the judicially-created doctrine of obviousness-type double patenting as being unpatentable over claim 73 of co-pending Application No. 09/249,543.

Claim 73 of Application No. 09/249,543 as entered requires in element (ii) forming a protein having an N-terminal cysteine or selenocysteine after intein mediated cleavage of a fusion protein. Claims 22 and 27-30 on the other hand, rely on ligating a chemically synthesized peptide or protein in which the synthesis includes creating an N-terminal cysteine. This is distinct from the claimed improvement in '543 where intein cleavage yields an N-terminal cysteine or selenocysteine in place of chemical synthesis claimed in the present application. There was no teaching in the art at the time of filing the above application that the improvement in '543 was possible or desirable.

Rejections under 35 U.S.C. §103

The Examiner has rejected the claims by combining Chong et al., Severinov et al. Muir et al., Telenti et al. and Comb et al. separately and together in various combinations with Burton et al and the Dictionary of Organic Chemical Compounds. All the references with the exception of Burton et al. have been separately discussed in the response dated 8/20/03 and rejections based on 35 U.S.C.§102 and 103 had been reversed. The issue at hand is the newly cited reference by Burton published in 1998 and the Dictionary of Organic Compounds which is undated.

Applicants respectfully request that the Examiner reverse the rejection for reasons that include (1) improper use of hindsight, (2) the absence of

any indication of motivation for combining references (3) secondary considerations of the widespread use of the claimed invention after its publication, (4) the teaching away from the claimed invention by Burton et al. ; and (5) the lack of motivation to select MESNA from a list of thiol reagents contained in the Dictionary of Organic Compounds.

Chong et al. describes purification of proteins not the intein-dependent joining of two proteins or a protein and a peptide. Moreover, Chong et al. neither suggests nor teaches the use of MESNA as a nucleophilic substitution reagent for intein mediated cleavage and protein ligation.

Severinov et al. describe protein ligation using the vectors supplied by New England Biolabs to express fusion proteins for purposes of studying transcription. (see Methods on page 16205). The reference teaches the use of thiophenol as a thiol reagent. It does not teach or suggest the use of 2-mercaptoethanesulfonic acid as a nucleophilic substitution reagent for intein mediated cleavage.

Muir et al. describe the use of New England Biolabs vectors containing inteins and a protein binding domain for preparing precursor proteins for cleavage. The reference utilizes thiophenol but neither suggests nor teaches the use of 2-mercaptoethanesulfonic acid as a nucleophilic substitution reagent for intein mediated.

Telenti et al. describe the Mxe GyrA intein, which is characterized in the reference with respect to splicing and cleavage but <u>not</u> with respect to its use in ligation. Moreover, there is no suggestion in Telenti et al. that

nucleophilic substitution in an intein mediated cleavage reaction could be effectively achieved with MESNA.

Comb et al describe intein mediated cleavage in the presence of DTT. There is no suggestion that 2-mercaptoethanesulfonic acid be used as a nucleophilic substitution reagent for intein mediated cleavage for protein ligation.

Burton et al.

A reference must be considered as a whole when making an obviousness rejection (MPEP 2141). When considered as a whole, the Burton reference is substantially different from the claimed invention. Without a teaching to do so, one of ordinary skill in the art would not refer to the Burton reference because it describes an entirely different chemistry and a different solution to an entirely different problem than described in the claimed invention. Even if the Burton reference were considered by one of ordinary skill in the art, no motivation is provided in the art to select MESNA in preference to other thiol reagents.

Differences between Burton et al. and the present claimed invention include:

(a) The Burton reference provides no motivation for selecting MESNA for any reason let alone intein mediated cleavage. Indeed the Burton reference teaches that any thiol reagent from a list of 15 (col 16, lines 1-8) would be equally effective for the purposes described in the reference. This list of thiol reagents includes MESNA and also other thiol reagents that the applicants have found to be unsuitable for nucleophilic substitution in intein mediated cleavage for protein ligation such as 3-mercaptopropionic acid.

This teaches away from the applicants' finding that certain thiol reagents in this list behave substantially differently from each other with respect to intein mediated cleavage reactions (Figure 2).

(b) The reference is directed exclusively to free radical chemistry. This is not the chemistry used in the claimed invention which relies on nucleophilic substitution for intein mediated cleavage and protein ligation. A protein chemist or person of ordinary skill in the art would recognize that a free radical reaction is different from a nucleophilic substitution reaction.

(c) The double bond, in Figure 3 of the Burton reference which is adjacent to a reactive carbon on an ethylene, must be broken to add the thiol reagent exemplified by MESNA. There is no such requirement in the present claimed invention.

(d) Burton et al. links the thiol reagent to ethylene to form a sorption substrate product. Target proteins that reversibly bind the thiol reagent are then purified chromatographically.

In the applicants' claimed invention, the thiol reagent is involved in nucleophilic substitution chemistry that results in the formation of a C-terminal thioester on a protein for covalent linkage to a peptide. The product of the reaction involving the nucleophilic substitution reaction is the covalently linked protein-peptide product and not as in Burton et al a protein or synthetic compound covalently linked to the thiol reagent.

The Examiner has asserted

> Burton et al. teach that any of several thiol reagents including MESNA will provide a reactive thiol ligand for the purpose of ligating a target compound such as a protein or peptide. See paragraph spanning column 5 and 6 and col 6, lines 41 and 42.

The paragraph preceding that cited by the Examiner in the Burton reference (col 6, lines 55-58) states

> Thiol containing ligands which are reactive with an ethynically unsaturated group pendent to the support matrix under free radical conditions are referred to as "reactive thiol containing ligands".

Previous to this paragraph in col 6, lines 37-38, the text adds to the definition of a thiol containing ligand, the requirement that the ligand,

> when bound to the support matrix is capable of binding target compounds during the sorption stage of chromatography

In Burton et al. target proteins are bound reversibly (including proteins as stated in col 6, lines 40-51) to the reactive thiol containing ligands during the sorption stage of chromatography.

In conclusion, applicants assert that the Burton reference teaches away from the present claimed invention alone or in combination with the other cited references and as such applicants respectfully request that the Examiner reverse the rejection.

<u>The Dictionary of Organic Compounds</u>

The references in column 1 of Burton et al. do not include the Dictionary identified by the Examiner. There is no specifically stated

motivation to combine the Dictionary with Burton et al.. Even if there was motivation, the combination would not describe the present claimed invention.

The Examiner has asserted those of ordinary skill in the art practicing intein mediated cleavage for protein ligation in the presence of the thiol reagent, DTT, would be motivated by Burton and the Dictionary of Organic Compounds to select MESNA. For the reasons given above, Applicants assert that Burton et al. teaches away from the claimed invention.

The issue at hand therefore is whether the Dictionary of Organic Compounds which describes multiple thiol reagents would be sufficient alone to motivate one of ordinary skill in the art to replace DTT with MESNA. This is a clear application of hindsight.

The Examiner has asserted that two criteria-pKa and solubility which are parameters contained in the Dictionary description of thiol reagents are sufficient to direct a skilled artisan to select the optimum thiol reagent for intein mediated cleavage suitable for peptide ligation. The Examiner has not however provided any teaching in the art to support this assertion.

In contrast, Applicants assert that it is unknown what role is played by the pKa of a thiol reagent in intein mediated cleavage. For example, DTT has two thiol groups having a pKa of 9.2 and 10.1, where this reagent is not a particularly effective thiol reagent for generating a protein having a C-terminal thioester suitable for protein ligation. The Examiner asserts that the higher the electronegativity of the thiol reagent the more suitable it is for intein mediated cleavage and peptide ligation. Not only is this unfounded,

it appears to be untrue also. In summary, there is no basis for the Examiner's assertion that a pKa of 9.7 alone renders a thiol reagent inferior to one with a pKa of 9.5 for intein mediated cleavage and protein ligation.

The Examiner has asserted that the second criterium for choosing MESNA is its solubility. However many of the thiol reagents tested by the applicants were soluble for example, DTT.

In the telephone interview, Examiner Nashed asserted a third criterium, namely that of smell where MESNA is not smelly and therefore is preferred. However, DTT is not smelly. Moreover, thiophenol which was a reagent of choice in native chemical ligation is exceedingly smelly.

Applicants respectfully assert that the Examiner's criteria for selecting a suitable thiol reagent for intein mediated cleavage is not supported by the art and is an exercise in hindsight.

In summary, the properties of a desirable thiol reagent for intein mediated cleavage for protein ligation do not rely solely on solubility or electronegativity. Other factors such as "leaving group" properties play a role. The ability of the thiol reagent to gain access to the intein active site also plays a role. Because the factors are complex and uncertain, applicants screened multiple agents before identifying MESNA.

Applicants assert that no motivation is provided for combining the references discussed above or for the rejection of the claims as obvious in light of such combination.

For the reasons provided above, applicants respectfully request the Examiner to remove the rejection.

CONCLUSION

For the reasons set forth above, Applicants respectfully submit that the rejections set forth in the Official Action of May 23, 2003 have been overcome and that this case is in condition for immediate allowance. Early and favorable consideration leading to prompt issuance of this Application is earnestly solicited. Applicants petition for an extension of three months under 37 C.F.R. 1.136 and enclose a check for $475 covering the extension fees. Applicants authorize that any additional fees that may be due be charged to deposit account number 14-0740.

Should the Examiner wish to discuss any of the amendments and/or remarks made herein, the undersigned Attorney would appreciate the opportunity to do so. Thus, the Examiner is hereby authorized to call the undersigned collect at the number shown below.

Respectfully submitted,

NEW ENGLAND BIOLABS, INC.

Harriet Strimpel

Date: May 28, 2004

Harriet M. Strimpel, D.Phil.
(Reg. No. 37,008)
Chief Patent Counsel
32 Tozer Road
Beverly, Massachusetts 01915
(978) 927-5054; Ext. 373

Reprinted from

GENE

AN INTERNATIONAL JOURNAL ON GENES AND GENOMES

Gene 192 (1997) 271–281

Single-column purification of free recombinant proteins using a self-cleavable affinity tag derived from a protein splicing element

Shaorong Chong [a], Fana B. Mersha [a], Donald G. Comb [a], Melissa E. Scott [a], David Landry [a], Luis M. Vence [a], Francine B. Perler [a], Jack Benner [a], Rebecca B. Kucera [a], Christine A. Hirvonen [a,1], John J. Pelletier [a], Henry Paulus [b], Ming-Qun Xu [a,*]

[a] *New England Biolabs, 32 Tozer Road, Beverly, MA 01915, USA*
[b] *Boston Biomedical Research Institute, Boston, MA 02114, USA*

ELSEVIER

SEVIER

Gene 192 (1997) 271–281

GENE
AN INTERNATIONAL JOURNAL ON GENES AND GENOMES

Single-column purification of free recombinant proteins using a self-cleavable affinity tag derived from a protein splicing element

Shaorong Chong [a], Fana B. Mersha [a], Donald G. Comb [a], Melissa E. Scott [a], David Landry [a], Luis M. Vence [a], Francine B. Perler [a], Jack Benner [a], Rebecca B. Kucera [a], Christine A. Hirvonen [a,1], John J. Pelletier [a], Henry Paulus [b], Ming-Qun Xu [a,*]

[a] *New England Biolabs, 32 Tozer Road, Beverly, MA 01915, USA*
[b] *Boston Biomedical Research Institute, Boston, MA 02114, USA*

Received 25 October 1996; received in revised form 24 January 1997; accepted 27 January 1997; Received by A. Bernardi

stract

A novel protein purification system has been developed which enables purification of free recombinant proteins in a single romatographic step. The system utilizes a modified protein splicing element (intein) from *Saccharomyces cerevisiae* (*Sce* VMA tein) in conjunction with a chitin-binding domain (CBD) from *Bacillus circulans* as an affinity tag. The concept is based on the oservation that the modified *Sce* VMA intein can be induced to undergo a self-cleavage reaction at its N-terminal peptide linkage y 1,4-dithiothreitol (DTT), β-mercaptoethanol (β-ME) or cysteine at low temperatures and over a broad pH range. A target rotein is cloned in-frame with the N-terminus of the intein-CBD fusion, and the stable fusion protein is purified by adsorption nto a chitin column. The immobilized fusion protein is then induced to undergo self-cleavage under mild conditions, resulting in ie release of the target protein while the intein-CBD fusion remains bound to the column. No exogenous proteolytic cleavage is eeded. Furthermore, using this procedure, the purified free target protein can be specifically labeled at its C-terminus. © 1997 lsevier Science B.V.

eywords: Protein purification; Intein; Chitin-binding domain; Protein labeling

l. Introduction

Purification of recombinant proteins by the use of affinity tags is a convenient and widely used technology (LaVallie and McCoy, 1995). A target protein is expressed as a fusion with an affinity tag which allows purification on an affinity column. A variety of affinity tags have been used, including *Schistosoma* glutathione *S*-transferase (GST) (Smith and Johnson, 1988), *Escherichia coli* maltose-binding protein (MBP) (Guan et al., 1988), *Staphylococcus* protein A (Nilsson and Abrahmsen, 1990), polyhistidine (Van Dyke et al., 1992), and calmodulin-binding peptide (Stofko-Hahn et al., 1992). Following protein purification, the affinity tag is typically cleaved from the target protein by treatment with a site-specific protease (LaVallie and McCoy, 1995). The use of proteases has limited the application of many affinity purification systems. First, cleavage by proteases is not always specific and may result in cleavage at secondary sites within the target protein. Second, the elevated temperatures required for many proteolytic cleavage reactions may unfavorably affect protein stability or activity. Third, the cleavage is sometimes inefficient due to the inaccessibility of the cleavage site on the fusion protein. Finally, additional chromatographic steps are required to separate the target protein from the affinity tag and the protease.

To avoid the limitations of using proteolytic cleavage, we have investigated the utilization of the inducible

* Corresponding author. Tel.: +1 508 9275054, ext. 372; fax: +1 508 9211350; e-mail: xum@neb.com
[1] Present address: Laboratory of Genetics, University of Wisconsin, Madison, WI 53706, USA.

Abbreviations: Amp^r, ampicillin resistance gene; β-ME, β-mercaptoethanol; bp, base pair(s); CBD, chitin-binding domain; DTT, 1,4-dithiothreitol; Intein, modified *Sce* VMA intein; IPTG, isopropyl-β-D-thiogalactopyranoside; kb, kilobase(s) or 1000 bp; kDa, kilodalton(s); *lacI*^{q}, gene overexpressing *lac* repressor; MBP or M, *E. coli* maltose-binding protein; MCS, multiple cloning site; NEB, New England Biolabs, Inc.; PAGE, polyacrylamide gel electrophoresis; P_{tac}, *tac* promoter; PCR, polymerase chain reaction; SDS, sodium dodecyl sulfate.

PII S0378-1119(97)00105-4

self-cleavage activity of a protein splicing element in protein purification (Fig. 1A). Protein splicing is a post-translational processing event in which an internal protein segment, the intein (Perler et al., 1994), can catalyze its own excision from a precursor protein and concomitantly ligate the flanking regions, the exteins, to form a mature protein (Kane et al., 1990; Hirata et al., 1990; Davis et al., 1992; Perler et al., 1992; Cooper et al., 1993; Xu et al., 1993). Using in vitro protein splicing systems, we have elucidated the protein splicing pathway of inteins from the thermostable DNA polymerase of *Pyrococcus* sp. GB-D and the 69-kDa vacuolar ATPase subunit of *Saccharomyces cerevisiae* (Xu et al., 1993, 1994; Shao et al., 1995, 1996; Chong et al., 1996; Xu and Perler, 1996). Through amino acid substitution, we have been able to modulate each step in the splicing pathway and convert the splicing reaction into an efficient polypeptide cleavage reaction at either splice junction (Xu and Perler, 1996; Chong et al., 1996). Specifically, a single amino acid substitution, Asn^{454} to Ala, at the C-terminal splice junction of the intein from the 69-kDa vacuolar ATPase subunit of *S. cerevisiae* (*Sce* VMA intein), has been shown to block splicing and C-terminal cleavage, but not the N-terminal cleavage mediated by the N-S acyl shift at Cys1 (Chong et al., 1996) (Fig. 1B). Nucleophiles that react with thioesters, such as thiols (i.e., β-ME, DTT, cysteine) and hydroxylamine, can effectively shift the N-S equilibrium by attacking the thioester, thereby inducing the N-terminal cleavage of the intein. In the case of cysteine-induced cleavage, a spontaneous S-N shift occurs after the nucleophilic attack by the sulfhydryl group of cysteine, resulting in formation of a native peptide bond between the cysteine and the C-terminus of the N-extein (Chong et al., 1996) (Fig. 1B).

Here we report the utilization of the inducible self-cleavage activity of the modified *Sce* VMA intein (Y) to develop a novel protein purification system. We demonstrate the effectiveness of the system by expressing a number of procaryotic and eucaryotic proteins in *E. coli* and purifying them free of their affinity tag after a single chitin column.

2. Results

2.1. Construction of the MYB fusion and characteristics of the CBD as an affinity tag

The maltose-binding protein from *E. coli* (MBP, or M) was fused to the N-terminus of the modified *Sce* VMA intein (Y) containing an Asn^{454}Ala substitution, and a small (5 kDa) chitin-binding domain (CBD or B) from the C-terminal region of chitinase A1 from *Bacillus circulans* WL-12 (Watanabe et al., 1994) was fused to the C-terminus of the intein. The resulting fusion construct, pMYB129, was used as a model system for studies of affinity binding and inducible cleavage at the N-terminal splice junction. The fusion protein MYB bound tightly to chitin beads. In a typical experiment, about 2 mg of the MYB fusion protein from crude cell extract bound to 1 ml of chitin beads (David Landry, NEB, unpublished results). High salt concentrations (up to 1 M NaCl), 0.1% Triton X-100 or Tween 20 and 1 mM EDTA had no effect on the binding of the CBD to chitin beads (data not shown). Buffers containing SDS (>0.5%) or guanidinium (6 M) were required to strip MYB or other CBD fusion proteins from chitin beads (data not shown).

2.2. Thiol-induced cleavage at the intein N-terminus in the MYB fusion protein

To study cleavage of the MYB fusion protein in solution, the unspliced MYB fusion protein (97 kDa) was purified on a single amylose column, with a yield of about 20 mg per liter of culture. The MYB fusion protein was very stable in vivo, during purification and during storage in solution at 4°C. Incubation of MYB at temperatures ranging from 25°C to 50°C resulted in no increase in M and YB (data not shown). The purified MYB fusion protein was treated at 4°C with different concentrations of DTT, β-ME and free cysteine (Fig. 2). More than 90% of MYB was cleaved by 30 mM DTT after 16 h incubation, generating M (42 kDa) and YB (55 kDa) (Fig. 2B), whereas β-ME was less effective (70% cleavage after 16 h; Fig. 2A). Cleavage by DTT was more efficient at high pH (Fig. 2D). Cleavage of MYB by cysteine was less efficient than cleavage by DTT or β-ME and required higher concentrations and higher pH values (Fig. 2C). The cleavage of MYB by hydroxylamine (100 mM) was also tested and found to be less efficient at 4°C than the thiol compounds (data not shown). In the absence of DTT, β-ME or cysteine, the MYB fusion protein was completely stable at 4°C in the 5.5–9.0 pH range. A typical control sample is shown in the first lane of each panel in Fig. 2. Furthermore, the cleavage reaction by DTT, β-ME or cysteine was efficient over a wide range of salt concentrations (50 mM–2 M NaCl) and in the presence of 0.1% Triton X-100 or Tween 20 (data not shown).

2.3. Cleavage of the MYB fusion protein immobilized on chitin beads

The crude cell extract containing MYB was mixed with chitin beads. After extensive wash to remove unbound proteins, different concentrations of β-ME or DTT were applied to the beads to induce the cleavage reaction at 4°C. Only MBP was eluted after a 6 h or 16 h incubation (Fig. 3A and B). These results indicate that cleavage occurred at the intein N-terminus while

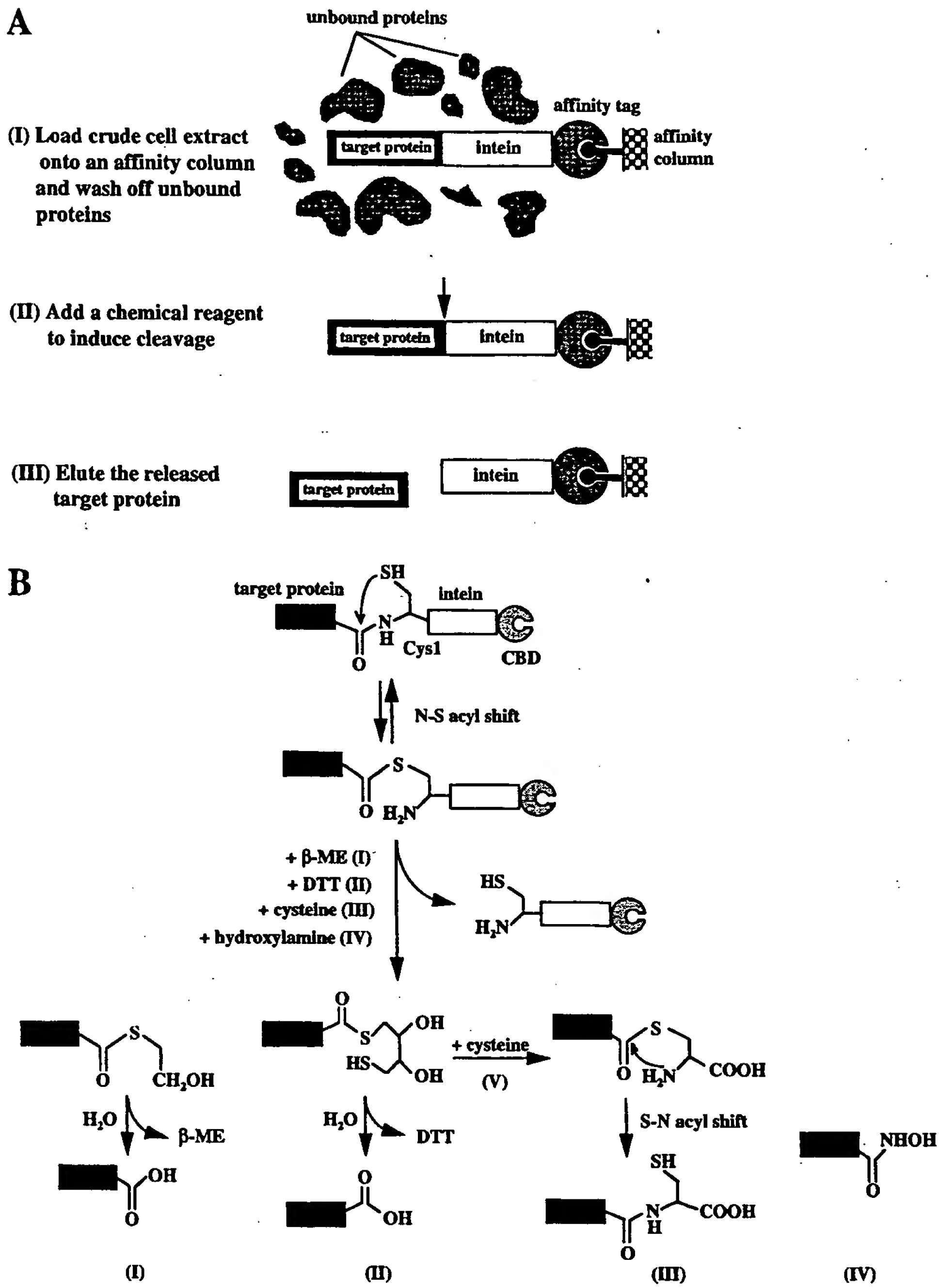


Fig. 1. (A) The concept of the intein-mediated protein purification with a self-cleavable affinity tag. A target protein is fused to the N-terminus of a modified protein splicing element (intein) which is in turn fused to an affinity tag (CBD). The fusion protein from a crude cell extract is purified by adsorption to a chitin affinity column (I). The intein is then induced to undergo on-column self-cleavage (arrow) by a chemical reagent (II). The target protein is specifically released from the column and eluted as a pure protein (III). (B) The chemical mechanism of the intein N-terminal cleavage reaction. The *Sce* VMA intein was modified by an $Asn^{454}Ala$ mutation which inactivates its splicing and C-terminal cleavage activities (Chong et al., 1996). This modified intein, when inserted between a target protein and a chitin-binding domain (CBD), catalyzes an N-S acyl shift at its N-terminal Cys1 which results in a thioester bond between the target protein and Cys1. Normally, the equilibrium of the N-S shift strongly favors peptide bond formation and the fusion precursor is quite stable. The presence of β-ME (I), DTT (II), cysteine (III) or hydroxylamine (IV) shifts the N-S equilibrium by cleaving the thioester bond. After cleavage, β-ME (I) or DTT (II) is hydrolyzed from the C-terminus of the target protein, while cysteine (III) or hydroxylamine (IV) remain covalently attached to the target protein. In the case of cysteine, a subsequent S-N shift results in formation of a native peptide bond. Cysteine can also attack the C-terminal thioester bond of the DTT-cleaved target protein and subsequently attach to the target protein through a peptide bond (V). Therefore, when a labeled cysteine is used, the target protein can be specifically labeled at its C-terminus.

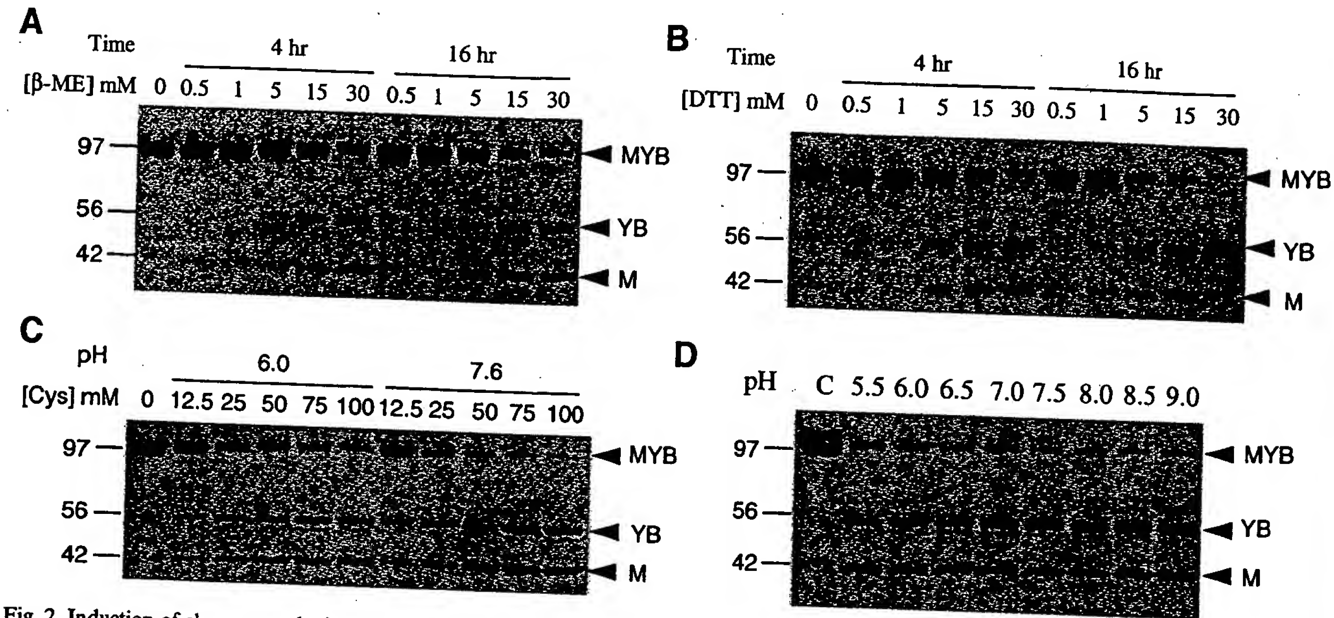


Fig. 2. Induction of cleavage at the intein N-terminus in the MYB fusion protein. Cleavage of purified MYB fusion protein (97 kDa) was induced with β-ME (A), DTT (B) or cysteine (Cys) (C) to generate the maltose-binding protein (M, 42 kDa) and the intein-CBD fusion (YB, 55 kDa). (D) pH effects on MYB cleavage by 30 mM DTT. Protein molecular weight standards (kDa, NEB) are shown on the left side of each panel. A control sample is shown in the first lane of each panel. In panel D, C is a control reaction at pH 8 without DTT. **Methods:** The gene encoding the *B. circulans* CBD was made synthetically using eight oligonucleotides (NEB) comprising both strands of the gene and overlapping by 12 bp. (1) 5′-T ATG ACG ACA AAT CCT GGT GTA TCC GCT TGG CAG GTC-3′, (2) 5′-pATA AGC TGT GTT GAC CTG CCA AGC GGA TAC ACC AGG ATT TGT CGT CA-3′, (3) 5′-pAAC ACA GCT TAT ACT GCG GGA CAA TTG GTC ACA TAT AAC GGC-3′, (4) 5′-pTTT ATA CGT CTT GCC GTT ATA TGT GAC CAA TTG TCC CGC AGT-3′, (5) 5′-pAAG ACG TAT AAA TGT TTG CAG CCC CAC ACC TCC TTG GCA GGA-3′, (6) 5′-pGGA TGG TTC CCA TCC TGC CAA GGA GGT GTG GGG CTG CAA ACA-3′, (7) 5′-pTGG GAA CCA TCC AAC GTT CCT GCC TTG TGG CAG CTT CAA TCG AGC T-3′, (8) 5′-CGA TTG AAG CTG CCA CAA GGC AGG AAC GTT-3′. The eight oligonucleotides were annealed and ligated into the *Nde*I and *Sac*I sites of plasmid pMIP22/23, a derivative of pMIP21 (Xu et al., 1993). The resulting plasmid (pBIP22/23) was used as a template for subsequent PCR amplification of the CBD gene. The PCR-synthesized CBD fragment was then used to construct pMIB115, which encodes the gene for a fusion protein consisting of the *E. coli* MBP, the *Psp* pol intein-1 (Xu et al., 1993) and the CBD (F.B. Mersha, unpublished). The modified *Sce* VMA intein (Y) containing an Asn[454]Ala substitution from pMYT1 (Chong et al., 1996) was inserted in-frame between M and B in pMIB115, replacing the *Psp* pol intein-1. The resulting vector, pMYB129, encodes a gene for the fusion protein MYB. The MYB fusion protein was expressed and purified on amylose resin as described by Chong et al. (1996). The purified MYB fusion protein (1.5 mg/ml) was incubated with various concentrations (0.5–30 mM) of β-ME or DTT in column buffer (20 mM Hepes, pH 8.0; 0.5 M NaCl) for 4 or 16 h at 4°C. For experiments on pH effects, the purified MYB was incubated at 4°C for 16 h with 30 mM DTT in the column buffer at different pH values (5.5–9.0). The N-terminal cleavage of MYB was also induced by treatment of MYB with various concentrations (12.5–100 mM) of cysteine at pH 6.0 or 7.6 in 20 mM Hepes, 0.5 M NaCl for 16 h at 4°C. As controls, the purified MYB was incubated at 4°C at pH 5.5–9.0 for 16 h in the absence of β-ME, DTT or cysteine. All reactions were stopped by mixing the protein samples with 1/2 vol. of 3× protein sample buffer (NEB) and boiling for 5 min. Each sample (2 μl) was electrophoresed on a 12% Tris-glycine PAGE (Novex). The running buffer was 25 mM Tris, pH 8.3, 190 mM glycine, 0.1% SDS. Staining was with Coomassie Brilliant Blue.

the MYB fusion protein was immobilized on the chitin beads, and that the cleavage conditions had no effect on the binding of the CBD. To examine the efficiency of on-column cleavage, the chitin beads were washed with 2% SDS to strip any remaining MYB fusion protein as well as the cleavage product (YB). More than 90% of the MYB fusion protein was cleaved by 30 mM DTT at 4°C in 16 h (data not shown).

2.4. Construction of the pCYB vectors for cloning of target proteins

The MBP sequence of pMYB129 was replaced with a series of multiple cloning sites, yielding the pCYB vectors (Fig. 4). The pCYB vectors have a ColE1 replication origin, an M13 origin for generating single-stranded DNA and an ampicillin resistance gene. Protein expression is under control of an IPTG-inducible *tac* promoter. The presence of *lac*I[q] repressor gene results in tighter transcriptional control. pCYB vectors with a T7 promoter were also constructed (data not shown). Three multiple cloning sites (MCS) of the pCYB vectors are shown in Fig. 4B. When the *Nde*I site is used for cloning a target gene, translation can initiate at the ATG of the *Nde*I site and the fusion protein will have the natural N-terminus of the target protein. The use of *Sap*I site of MCS1 will result in the fusion of the C-terminus of the target protein to the N-terminal cysteine of the intein (Cys1). Therefore, when the *Nde*I and *Sap*I sites of MCS1 are used, the native sequence of the target

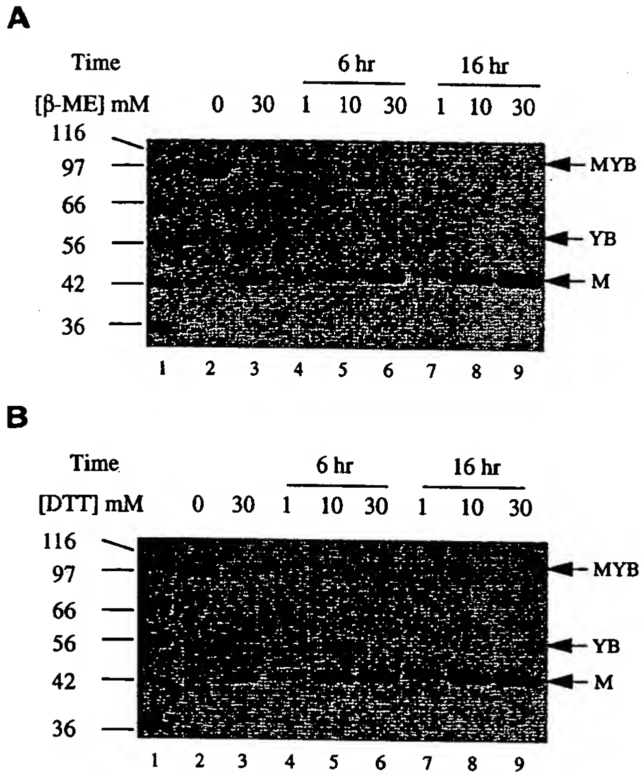
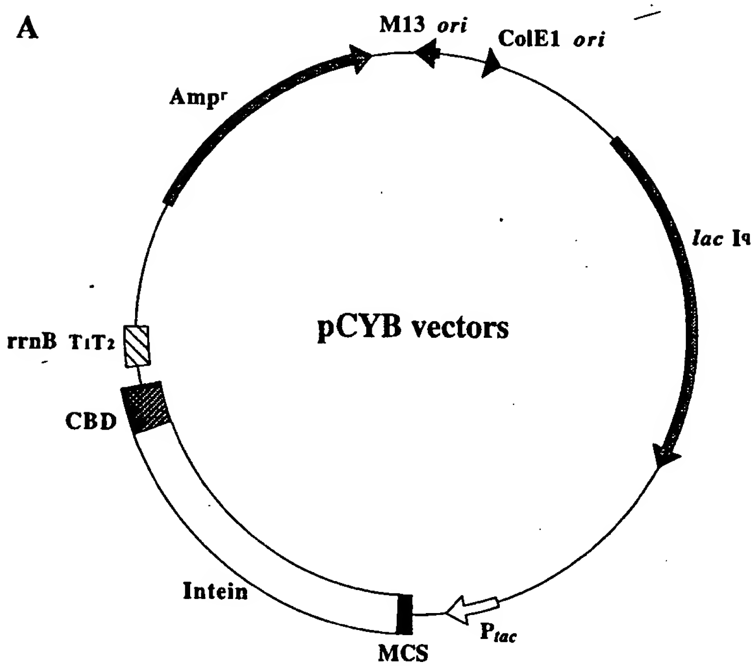


Fig. 3. Induction of cleavage of MYB immobilized on chitin beads with β-ME (A) or DTT (B). Lanes: 1, protein molecular weight standards (kDa, NEB); 2, purified MYB; 3, cleavage of the purified MYB in solution by 30 mM β-ME or DTT at room temperature for 6 h; 4–9, eluate from chitin beads after treatment of bound MYB protein with different concentrations of β-ME or DTT at 4°C for 6 or 16 h. The samples were examined by Coomassie Blue staining. M, maltose-binding protein; YB, intein-CBD fusion. **Methods:** Spherical chitin beads were prepared by emulsification of a solution of chitosan in 10% acetic acid with an organic solvent containing a suitable surfactant, followed by the addition of acetic anhydride (Roberts and Taylor, 1988). After gelation, the resulting beads (50–100 mm) were washed with water. Following dehydration by washing with methanol, the beads were reacetylated with acetic anhydride. The degree of *N*-acetylation was determined with the 2,4,6-trinitrobenzenesulfonic acid and perchloric acid assay using a glucosamine standard (Wilkie and Landry, 1988). Chitin beads are commercially available from NEB (cat. No. 6651). The MYB fusion protein was expressed as described by Chong et al. (1996). 1 ml of the clarified cell extract was mixed with 0.2 ml chitin beads in an Eppendorf tube. After 10 min incubation at 4°C, the beads were washed 5 times with column buffer. β-ME or DTT at various concentrations (1–30 mM) in the column buffer was then applied to the beads and the samples were incubated at 4°C for 6 or 16 h. Chitin beads were spun down and the supernatant was collected and analyzed by SDS-PAGE as described in Fig. 2.

protein will be obtained after cleavage, with no vector-derived amino acid residues. When a target gene is cloned into the *Nde*I and *Sma*I sites of MCS2, the target protein will have an additional glycine residue at its C-terminus after cleavage. When the *Nde*I and *Xho*I sites of MCS3 are used, the target protein will have three additional residues (leucine-glutamate-glycine) at its C-terminus.

2.5. Expression and purification of target proteins using pCYB vectors

To demonstrate that the pCYB vectors can be used to express and purify recombinant proteins, a number of proteins from both procaryotes and eucaryotes were cloned into the pCYB vectors and purified after a single chitin column as summarized in Table 1. The procedures for the expression and purification of *Hha*I methylase (Caserta et al., 1987) are illustrated in Fig. 5A. After induction with IPTG, the fusion protein (approx. 95 kDa) could be observed in the crude extract (Fig. 5A, lane 3), but not in uninduced cells (Fig. 5A, lane 2). After the crude extract from the induced cells was loaded onto a chitin column, most of the fusion protein bound to the chitin resin (Fig. 5A, lane 4). The column was then washed with column buffer. DTT (30 mM) was then quickly passed through the column to induce the on-column cleavage reaction. No significant amount of *Hha*I methylase was cleaved off the column during this brief DTT flush (Fig. 5A, lane 5). The column was then incubated at 4°C overnight, followed by elution of released *Hha*I methylase using the column buffer without DTT (Fig. 5A, lanes 6–9). The enzyme was more than 95% pure as estimated by staining an SDS-PAGE gel with Coomassie Blue and was active as determined by the standard methylase assay (Caserta et al., 1987) (data not shown). Washing the column with 2% SDS to strip the bound protein indicated that more than 90% of the fusion protein was cleaved (Fig. 5A, lane 10). The expression and purification of some of the recombinant proteins are shown in Fig. 5B. The specific activities of the purified *Bst* DNA polymerase large fragment (Aliotta et al., 1996) and the *Bam*HI, *Bgl*II, and *Eco*RV restriction endonucleases (Brooks et al., 1989; Kravets et al., 1990) were determined and found to be more than 70% of the values obtained with enzymes purified by conventional methods (Table 1). The yield of the *Bst* DNA polymerase large fragment was increased by 4-fold if a T7 promoter was used in place of a *tac* promoter. The N-terminal catalytic domain (residues 1–177) of *Brugia malayi* cyclophilin PPIase A (Page et al., 1995) was purified and a minor band of approx. 55 kDa corresponding to the intein-CBD fusion partner was also present in the eluate. The purified calmodulin-dependent protein kinase (CamKII) (Lin et al., 1987) showed an calmodulin-dependent protein kinase activity similar to the wild-type CamKII purified by a eucaryotic expression system, albeit with reduced activity (S.T. Wong-Madden, NEB, personal communication). The CamKII sample also contained a minor species migrating slightly slower than the major CamKII band. The purified T4 DNA ligase exhibited 90% of the specific activity of the enzyme purified by multiple chromatographic columns (Charles Card, NEB, personal communication).



B

```
        Met Ala Ser Ser Arg Val Asp Ser Arg Glu Phe Leu Glu Gly Ser Ser Cys1
MCS1 CAT ATG GCT AGC TCG CGA GTC GAC TCT AGA GAA TTC CTC GAG GGC TCT TCC TGC
      NdeI   NheI    NruI    SalI    XbaI    EcoRI   XhoI      SapI

        Met Ala Ser Ser Arg Val Asp Ser Arg Glu Phe Leu Glu Pro Gly Cys1
MCS2 CAT ATG GCT AGC TCG CGA GTC GAC TCT AGA GAA TTC CTC GAG CCC GGG TGC
      NdeI   NheI    NruI    SalI    XbaI    EcoRI   XhoI    SmaI

        Met Ala Ser Ser Arg Val Asp Ser Arg Glu Phe Leu Glu Gly Cys1
MCS3 CAT ATG GCT AGC TCG CGA GTC GAC TCT AGA GAA TTC CTC GAG GGG TGC
      NdeI   NheI    NruI    SalI    XbaI    EcoRI   XhoI
```

Fig. 4. (A) The pCYB *E. coli* expression vectors. All pCYB vectors contain the modified *Sce* VMA intein (Y, 1362 bp), carrying an $Asn^{454}Ala$ substitution, and the *B. circulans* chitin-binding domain (CBD or B, 135 bp). The multiple cloning sequence (MCS) was designed for insertion of the target gene in-frame with the downstream *Sce* VMA intein-CBD coding region. Expression of intein-CBD fusion proteins is under control of a *tac* promoter and regulated by the *lacI*^{q} repressor. The rrnB transcription terminators ($rrnBT_1T_2$) are located downstream from CBD. The pCYB vectors have a ColE1 replication origin, an M13 origin for generating single-stranded DNA and an ampicillin resistance gene. (B) Multiple cloning sites (MCS) of the pCYB vectors. Translation initiates at the ATG of the *Nde*I site. The amino acid sequence is shown above the nucleotide sequence and restriction enzyme sites are indicated below the nucleotide sequence. Cys1, the first (N-terminal) amino acid residue of the intein and the site of cleavage, is not part of the multiple cloning site. The recognition sequences of *Nde*I, *Sap*I, *Sma*I and *Xho*I are underlined. The ATG sequence of the *Nde*I site in pMIB124, a derivative of pMIB115, is 9 bp downstream from the ribosome-binding site of the *E. coli* malE gene (Guan et al., 1988) and is used as a translation start codon. **Methods:** pMIB124 was constructed by replacing the *E. coli* MBP coding sequence in pMIB115 with a polylinker containing *Nde*I and *Xho*I sites. pMIB124 was digested by *Xho*I and *Pst*I and ligated with the 1.5 kb *Xho*I-*Pst*I fragment containing the *Sce* VMA intein-CBD coding region from pMYB129 replacing the *Psp* pol intein-1-CBD sequence. The resulting plasmid pMYB131 was digested by *Nde*I-*Xho*I and ligated with a linker containing MCS3 to yield pCYB138 (6840 bp). pCYB1 and pCYB2 vectors, with MCS1 and MCS2 respectively, were created by introducing *Sap*I and *Sma*I site, respectively, through linker substitution of the region (flanking the N-terminus of the intein) between *Xho*I and *Kpn*I sites of pCYB138.

2.6. *In vitro labeling of MBP and* Bam*HI*

Following DTT-induced on-column cleavage, purified proteins were tested for replacement of the C-terminal DTT moiety by radiolabeled cysteine. The initial nucleophilic attack by the sulfhydryl group of cysteine followed by a spontaneous rearrangement (S-N shift) results in a peptide bond between the cysteine and the C-terminus of the target protein (Chong et al., 1996; Dawson et al., 1994) (Fig. 1B). Aliquots of MBP and *Bam*HI eluted from the chitin beads by DTT-induced cleavage were incubated with L-[^{35}S]cysteine (at 1 μM final concentration) immediately after elution. As shown in Fig. 5B, MBP (MBP panel: lane 5) and *Bam*HI (*Bam*H panel: lane 4) were both labeled. On the other hand, when a sample of MBP, purified on amylose resin (NEB) by elution with maltose, was incubated with L-[^{35}S] cysteine, no labeling was observed (MBP panel: lane 4).

Table 1
Recombinant proteins expressed and purified by pCYB vectors

Target protein[a]	Molecular weight (kDa)	Cloning sites	Promoter	Host *E. coli* strain[b]	Final yield (mg/l culture)	Activity[c]
*Hha*I methylase	40	*Nde*I, *Xho*I (MCS3)	*tac*	ER2267	3	+
MBP	42	*Nde*I, *Xho*I (MCS3)	*tac*	ER2267	20	n.d.
Bst DNA polymerase large fragment	65	*Nde*I, *Xho*I (MCS3)	T7	ER2504	2	90%
			tac	ER2504	0.5	+
T4 DNA ligase	60	*Nde*I, *Sma*I (MCS2)	*tac*	ER2267	4	90%
*Bam*HI[d]	26	*Nde*I, *Sap*I (MCS1)	T7	ER2504	1	75%
*Bgl*II[d]	24	*Xba*I, *Xho*I	T7	ER2504	4	99%
*Eco*RV[d]	26	*Nde*I, *Xho*I (MCS3)	T7	ER2504	0.5–1	72%
Cyclophilin PPIase	23	*Nde*I, *Xho*I (MCS3)	*tac*	ER2267	0.5	+
CamKII	33	*Nde*I, *Sap*I (MCS1)	T7	ER2504	0.8	+
			tac	ER2504	0.8	+

[a]Abbreviations are as in the legend to Fig. 5. The methods for cloning are described in Fig. 5.
[b]ER2267 (Fomenkov et al., 1994) is F^- $e14^-$ ($McrA^-$) *end* A1 SupE44 *thi*-1 Δ(malB) Δ(*argF-lac*)U169 Pro^+zjc::Tn*5*(Kan^r) *fhu*A2. ER2504 is a derivative of BL21(DE3) (Studier et al., 1990) (W. Jack and E.A. Raleigh, NEB, unpublished).
[c]The specific activity of each enzyme determined by an appropriate assay is shown as the percentage of the specific activity of the conventionally purified enzyme. In cases where the specific activities were not measured quantitatively, a plus (+) indicates that the enzyme showed activity after purification. n.d., the activity was not determined.
[d]The toxicity of even low levels of restriction endonuclease required the use of pCYB vectors with a T7 promoter and the co-expression of a corresponding methylase gene.

3. Discussion

In this paper, we describe a novel protein purification system which utilizes the controlled N-terminal cleavage activity of a modified protein splicing element. Unlike other protein purification methods employing affinity tags, this system allows a single-column purification of free recombinant proteins from crude cell lysates through a unique on-column cleavage reaction induced by simple and inexpensive chemical reagents. The intein-mediated cleavage separates the target protein from the affinity tag and eliminates the need for exogenous proteolytic cleavage and chromatographic removal of the protease and affinity tag from the target protein.

The key component of this system is the *Sce* VMA intein modified to undergo cleavage reaction only at its N-terminus. The MYB fusion protein was employed as a model system to study the cleavage reaction. A unique feature of this modified *Sce* VMA intein is that the cleavage reaction can be induced at 4°C (a preferred temperature for purification of proteins) (Fig. 2). Basic pH values favored DTT-induced cleavage, probably owing to the fact that the cleavage reaction is a nucleophilic substitution involving thiolate anion as the nucleophile. However, after 16 h incubation at 4°C, cleavage was mostly complete at all pH values ranging from 5.5 to 9.0 (Fig. 2D). This can be an advantage for purifying proteins with different pH-stability profiles. Furthermore, the ability to carry out cleavage reactions at high salt concentrations or in the presence of non-ionic detergents may improve the solubility of some target proteins. β-ME and DTT at concentrations up to 5 mM did not result in significant cleavage of MYB protein (Fig. 2), and target proteins that are sensitive to oxidation can therefore be purified in the presence of low concentrations of thiol reagents. After inducing cleavage, both β-ME and DTT formed a thioester bond with the C-terminus of the target protein (Fig. 1B). In the case of the MYB fusion protein, more than 90% of the β-ME or DTT moiety was hydrolyzed from the eluted MBP after 1 week storage at 4°C, pH 8.0, resulting in an unmodified MBP with a native C-terminus (data not shown). The intein-mediated cleavage reaction is highly specific and the target protein should therefore have a homogeneous C-terminus. This was supported by mass spectrometric analysis and sequencing of the C-terminus of the MBP domain after thiol-induced cleavage (Chong et al., 1996). Cysteine or hydroxylamine can also induce N-terminal cleavage reaction, resulting in a target protein with a modified C-terminus (Fig. 1B). For target proteins sensitive to reducing reagents (e.g., owing to the presence of intramolecular disulfide bonds), hydroxylamine can be used for inducing cleavage instead of β-ME and DTT.

To facilitate the purification of fusion proteins, the chitin-binding domain (CBD) from *B. circulans* was incorporated into the system. The CBD has an extremely high affinity for the chitin beads, which allows better recovery of the fusion protein from the crude cell extract. In addition, stringent wash conditions (e.g., high salt concentration or use of detergents) can be employed to reduce non-specific binding, thus increasing purity.

Another unique feature of this system is its ability to allow specific labeling of the C-terminus of a target protein. Since the labeled cysteine only reacts with the C-terminal DTT moiety of the purified target protein,

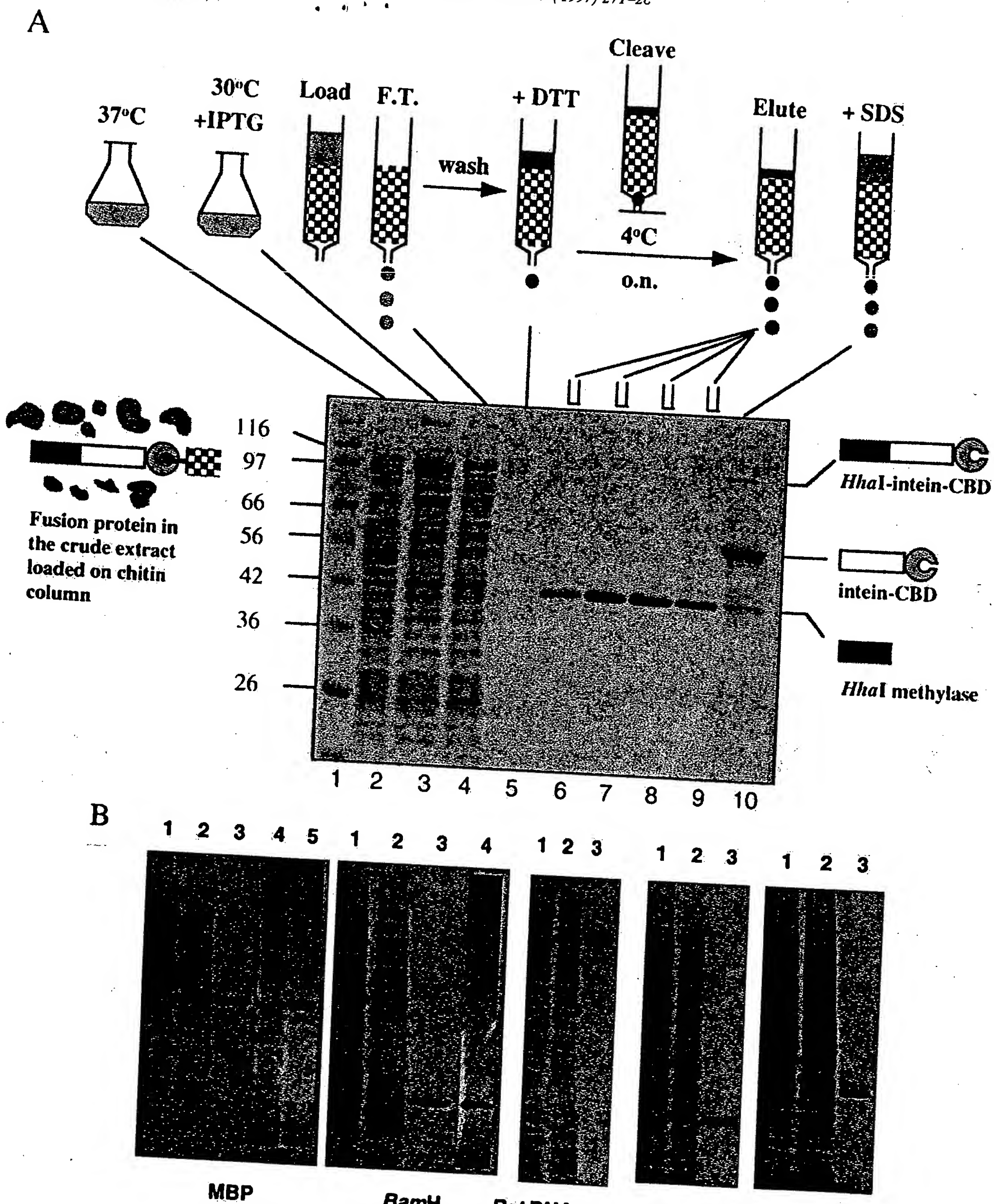


Fig. 5. Single column purification of recombinant proteins from *E. coli* using pCYB vectors. (A) Schematic representation of the expression and purification of *Hha*I methylase using a pCYB vector. Samples taken from different steps during the expression and purification procedures were separated by SDS-PAGE and the gel was stained with Coomassie Blue. Lanes: 1, protein molecular weight standards (kDa, NEB); 2, uninduced cell extract; 3, induced cell extract; 4, flow through (F.T.) from the load. After loading, the column is washed with column buffer until the protein content of the eluate reached a minimum; 5, flow through from the quick DTT flush; 6–9, the first four fractions of the elution after the 4°C overnight (o.n.) incubation in the presence of DTT; 10, a fraction from the SDS elution. (B) Purified proteins after a single chitin column. MBP, *E. coli* maltose-binding protein (42 kDa); *Bam*HI, a *Bacillus amyloliquefaciens* H restriction endonuclease (26 kDa); *Bst* DNA pol, *Bacillus stearother*

this method of protein labeling is more specific, less laborious and less likely to affect the structure and activity of the target protein than other methods of in vitro protein labeling (Shumacher and Tsomides, 1996). The utilization of purified proteins labeled in vitro at their C-termini may facilitate the investigation of their biological functions by newly developing techniques such as protein footprinting (Shumacher and Tsomides, 1996). In addition, this system provides an attractive alternative method for peptide synthesis and purification. Furthermore, the thiol esters that result from intein-mediated cleavage induced by thiol compounds can serve as intermediates in peptide ligation (Dawson et al., 1994).

Like other commonly used expression vectors, the pCYB vectors use a *tac* or T7 promoter for expression of heterologous proteins in *E. coli*. Other protein fusion systems can sometimes increase the expression of the fusion protein by placing a highly expressed affinity tag at the N-terminus of the fusion (LaVallie and McCoy, 1995; Smith and Johnson, 1988; Guan et al., 1988) (e.g., MBP, GST, etc.). In the pCYB vectors, however, translation of the fusion protein starts with the target protein, and therefore the expression level of the fusion protein is primarily determined by the expression of the target protein itself. We have found that the level of protein expression in the pCYB vectors, like T7 promoter vectors, varies with different target proteins. Using a T7 promoter often improves the expression of poorly expressed proteins (Studier et al., 1990). However, some target proteins, especially those from eucaryotes, are expressed poorly in the pCYB vectors (data not shown). In the case of CamKII, the purified protein also had a lower activity compared to the enzyme purified by a

mophilus DNA polymerase large fragment (68 kDa); Cyclophilin, *Brugia malayi* cyclophilin PPIase N-terminal domain (23 kDa); CamKII, a calmodulin-dependent protein kinase (33 kDa). For each panel, lanes: 1, crude cell extract; 2, flow through; 3, a fraction after DTT-induced cleavage; for MBP, 4, control sample for in vitro labeling using amylose-purified MBP; 5, in vitro labeling of chitin-purified MBP with L-[^{35}S]cysteine; for *Bam*HI, lane 4, in vitro labeling with L-[^{35}S]cysteine. **Methods:** All target genes were amplified by PCR to include *Nde*I site at their N termini. The gene for *Hha*I methylase was amplified by PCR from pNW2801 (Caserta et al., 1987) (a gift of Dr. Geoffrey Wilson, NEB). The forward primer, 5′-GGC GGC <u>CAT ATG</u> **ATT GAA ATA AAA GAT AAA CAG** CTC-3′, contains an *Nde*I site (underlined) where the translation initiates at the ATG codon, and the first eight codons of the *Hha*I methylase gene (bold). The reverse primer, 5′-GGC GGC <u>CTC GAG</u> **ATA TGG TTT GAA ATT TAA TGA TGA** ACC-3′, contains an *Xho*I site (underlined) and the sequence of the last 9 amino acids of the *Hha*I methylase gene (bold). PCR mixtures (100 µl) contained Vent DNA polymerase buffer (NEB), 4 mM magnesium sulfate, 400 µM of each dNTP, 1 µM of each primer, 50 ng pNW2801 DNA and 1.0 unit of Vent DNA polymerase (NEB). Amplification was carried out using a Perkin-Elmer/Cetus thermal cycler at 94°C for 1 min, 50°C for 1 min and 72°C for 2 min, for 22 cycles. The PCR products were gel-purified and, after double digestion with *Nde*I and *Xho*I, directly ligated to a gel-purified *Nde*I-*Xho*I double-digested pCYB vector carrying a *tac* promoter and MCS3 to yield pCYB(MHha). *E. coli* strain ER2267 was transformed by pCYB(MHha) and the cells were grown and induced as described below for pMYB129. Similar procedures were followed in the cloning and expression of proteins shown in B and Table 1. For MBP, the expression vector was pMYB129. A 1-l culture was induced at 30°C by 0.3 mM IPTG for 4 h. For *Bst* DNA polymerase large fragment, the cell culture (1 l) was induced at 30°C for 4 h. For *Bam*HI, a 100-l culture was grown in a New Brunswick Magnaferm fermentor at 37°C and then induced at 30°C for 4 h. The pCYB vectors with a T7 promoter contain an *Xba*I site between the promoter and the ribosome-binding site. A PCR-generated fragment containing the gene for *Bgl*II (Anton et al., 1997), the ribosome-binding site and translation start codon (ATG) was cloned into the *Xba*I and *Xho*I sites of a pCYB vector with a T7 promoter, and the cells (1 l) were induced at 15°C for 16 h. For *Eco*RV, the culture (1 l) was induced at 15°C for 16 h. Expression of restriction endonucleases *Bam*HI (Brooks et al., 1989), *Bgl*II or *Eco*RV (Kravets et al., 1990) requires the use of pCYB vectors with a T7 promoter and the co-expression of a corresponding methylase gene in order to prevent the endonuclease from being too toxic to the cells. In the case of *Bam*HI or *Bgl*II, a compatible plasmid (a pACYC184 derivative, NEB) containing the corresponding methylase gene was co-transformed into the cells. In the case of *Eco*RV, the methylase gene was cloned into the same vector as the endonuclease gene but under its own promoter. The coding sequence of the N-terminal catalytic domain (residues 1–177) of *Br. malayi* cyclophilin PPIase was amplified from pMAL-CYP10 (a gift of Clotilde Carlow, NEB) and cloned into pCYB138, yielding pCYP127. The cells harboring pCYP127 (1 l) were induced at 15°C for 16 h. The gene for a calmodulin-dependent protein kinase (CamKII) was amplified by PCR from pSRalpha-CaMK (a gift of Dr. Howard Shulman), and the cells (1 l) were induced at 15°C for 16 h. The coding sequence of T4 DNA ligase (Armstrong et al., 1983) was amplified from a wild-type phage T4D strain (a gift of Dr. David Shub) and cloned into MCS2 (Table 1), yielding pT4YB1. ER2267 cells bearing pT4YB1 were induced by 0.3 mM IPTG at 15°C for 16 h. All purification procedures were conducted at 0–4°C. The cell pellet (3–5 g) from each 1-l culture was resuspended in 50 ml of column buffer (20 mM Hepes, pH 8.0; 0.5 M NaCl) and broken by sonication. After centrifugation at $25\,000 \times g$ for 30 min, the supernatant was passed through a column packed with 20 ml of chitin beads (NEB) at a flow rate of 0.5 ml/min and the column was then washed with the column buffer at a flow rate of 2 ml/min until the protein content of the eluate reached a minimum (OD_{595nm} <0.05 as measured by Bradford assay). Column buffer (50 ml) containing 30 mM DTT was quickly passed through the column in order to distribute DTT evenly throughout the resin and the flow was stopped. The column was incubated at 4°C overnight. Fractions (3 ml each) containing the target protein were obtained by eluting the column with the column buffer in the absence of DTT. After elution, 50 ml of 2% SDS in the column buffer were passed through the column at room temperature to strip the remaining bound proteins and fractions (4 ml each) were collected. The samples were analyzed by SDS-PAGE as described in Fig. 2. The results for *Hha*I methylase are shown in A. The expression and purification of *Bam*HI was similar but conducted on a production scale. The cell pellet (120 g) was resuspended in 500 ml column buffer before sonication and the crude cell extract was passed through a 180 ml column of chitin beads (NEB). The DTT-induced cleavage was conducted at 4°C for 1–3 days. For in vitro labeling of proteins, MBP (1 mg/ml) or *Bam*HI (0.5 mg/ml) was obtained by induction of on-column cleavage of the fusion proteins as described above. An aliquot (36 µl) of each freshly eluted sample was mixed with 4 µl of L-[^{35}S]cysteine (11.0 mCi/ml, 0.0102 µmol/ml, Du Pont NEN) and incubated at 4°C overnight. In a control reaction, an aliquot (36 µl) of MBP (1 mg/ml) purified on an amylose resin column (NEB) was incubated with 4 µl of L-[^{35}S]cysteine. Sample were electrophoresed on a 12% Tris-glycine PAGE (Novex) visualized by autoradiography.

eucaryotic system. This could be due to protein misfolding or lack of post-translational modifications in *E. coli*, or sensitivity to the high concentration of thiol reagent during purification. The development of an analogous eucaryotic expression system may help to improve the expression and activities of eucaryotic target proteins.

The level of protein expression in the pCYB vectors was also affected by induction temperature. We found that in most cases (e.g., *Bst* pol, *Bam*HI, *Bgl*II, *Eco*RV), fusion protein expression was higher at lower induction temperatures (15–30°C) than at 37°C (data not shown). However, in some cases (e.g., MBP), the induction temperature made no difference in expression. Thus, different induction conditions should be tested for each target protein in order to achieve optimal induction. Due to the specificity of the cleavage reaction and high affinity of CBD for chitin beads, pCYB vectors usually resulted in highly purified target proteins. However, when the expression of the target protein was inefficient, significant amounts of other proteins were co-purified (Fig. 5B). In the case of cyclophilin and CamKII, whose expression and final yield were relatively low (Table 1), some impurities could be seen on SDS-PAGE, probably as a result of protease degradation, unspecific binding to chitin beads, or elution of the fusion protein or fusion partner (i.e., intein-CBD fusion).

pCYB vectors are N-terminal fusion vectors in which the C-terminus of a target protein is fused to the N-terminus of the intein. The structure or sequence at the C-terminus of the target protein may affect cleavage efficiency. Preliminary data indicate that some residues, when placed immediately adjacent to Cys1, can either block the cleavage reaction or cause in vivo cleavage (data not shown). Therefore, in some cases, the addition of a favorable residue or sequence between the target protein and the intein may be necessary for the controllable cleavage reaction (e.g., using pCYB vectors with MCS2 and MCS3 for cloning). Furthermore, incomplete processing of N-terminal methionine in some fusion proteins may generate a target protein with heterogeneous N-termini. Nevertheless, the method described here clearly has many advantages and has been shown to drastically simplify protein purification. Using the pCYB vectors, many proteins were successfully purified with just one column with a yield comparable to or greater than the yield using conventional methods. Furthermore, most of the tested target proteins retain their activities after purification (Table 1).

In conclusion, by combining a modified protein splicing element with an affinity tag in protein purification, we have been able to obtain highly purified non-tagged recombinant proteins in a single chromatographic step. Owing to the self-catalytic property of the intein, this purification system does not require any cellular co-factor for cleavage, and the core components of this system (i.e., intein-CBD) should therefore be compatible with other expression and purification systems (e.g., yeast, baculovirus, mammalian cells, etc.).

Acknowledgement

We thank Ira Schildkraut, Richard Roberts, Yang Shao, William Jack, Christopher Noren, Paul Riggs, Elisabeth Raleigh, Maurice Southworth, Yi Tan, Michael Comb, Daniel Panne and Jill Platko for valuable discussions and reading of the manuscript; Geoffrey Wilson for the gift of pNW2801 containing *Hha*I methylase gene; William Jack, Richard Morgan, Jurate Bitinaite for the clones of restriction endonucleases and methylases; Gloria-Jane Healy, Richard Whitaker, Sharon Wong-Madden, Diana T. O'Loane, Dong Ma, Tao Cui, Robyn Byer and Charles Card for technical assistance; David Shub for providing phage T4D strain; Howard Shulman for plasmid pSRalpha-CaMK; Huimin Kong for the *Bst* DNA polymerase large fragment clone; and Clotilde Carlow for the clone of the *Br. malayi* cyclophilin PPIase.

References

Aliotta, J.M., Pelletier, J.J., Ware, J.L., Moran, L.S., Benner, J.S., Kong, H., 1996. Thermostable *Bst* DNA polymerase I lacks a 3′→5′ proofreading exonuclease activity. Genet. Anal. Biomol. Eng. 12, 185–195.

Anton, B.P., Heiter, D.F., Benner, J.S., Hess, E.J., Greenough, L., Moran, L.S., Slatko, B.E., Brooks, J.E., 1997. Cloning and characterization of the *Bgl*II restriction-modification system reveals a possible evolutionary footprint. Gene 187, 19–27.

Armstrong, A., Brown, R.S., Tsugita, A., 1983. Primary structure and genetic organization of phage T4 DNA ligase. Nucleic Acids Res. 11, 7145–7156.

Brooks, J.E., Benner, J.S., Heiter, D.F., Silber, K.R., Sznyter, L.A., Jager-Quinton, T., Moran, L.S., Slatko, B.E., Wilson, G.G., Nwankwo, D.O., 1989. Cloning the *BamHI* restriction-modification system. Nucleic Acids Res. 13, 979–997.

Caserta, M., Zacharias, W., Nwankwo, D., Wilson, G.G., Wells, R.D., 1987. Cloning, sequencing, in vivo promoter mapping and expression in *Escherichia coli* of the gene for the *Hha*I methyltransferase. J. Biol. Chem. 262, 4770–4777.

Chong, S., Yang, S., Paulus, H., Benner, J., Perler, F.B., Xu, M.-Q., Protein splicing involving the *Saccharomyces cerevisiae* VMA intein: the steps in the splicing pathway, side reactions leading to protein cleavage, and establishment of an in vitro splicing system. 1996. J. Biol. Chem. 271, 22159–22168.

Cooper, A.A., Chen, Y., Lindorfer, M.A., Stevens, T.H., 1993. Protein splicing of the yeast TFP1 intervening protein sequence: a model for self-excision. EMBO J. 12, 2575–2583.

Davis, E.O., Jenner, P.J., Brooks, P.C., Colston, M.J., Sedgwick, S.G., 1992. Protein splicing in the maturation of *M. tuberculosis* RecA protein: a mechanism for tolerating a novel class of intervening sequence. Cell 71, 201–210.

Dawson, P.E., Muir, T.W., Clark-Lewis, I., Kent, S.B.H., 1994. Synthesis of proteins by native chemical ligation. Science 266, 776–779.

Fomenkov, A., Xiao, J.-P., Dila, D., Raleigh, E., Xu, S.-Y., 1994. The

'endo-blue method' for direct cloning of restriction endonuclease genes in *E. coli*. Nucleic Acids Res. 22, 2399–2403.

Guan, C. di, Li, P., Riggs, P.D., Inouye, H., 1988. Vectors that facilitate the expression and purification of foreign peptides in *Escherichia coli* by fusion to maltose-binding protein. Gene 67, 21–30.

Hirata, R., Ohsumi, Y., Nakano, A., Kawasaki, H., Suzuki, K., Anraku, Y., 1990. Molecular structure of a gene, VMA1, encoding the catalytic subunit of H^+-translocating adenosine triphosphatase from vacuolar membranes of *Saccharomyces cerevisiae*. J. Biol. Chem. 265, 6726–6733.

Kane, P.M., Yamashiro, C.T., Wolczyk, D.F., Neff, N., Goebl, M., Stevens, T.H., 1990. Protein splicing converts the yeast TFP1 gene product to the 69-kD subunit of the vacuolar H^+-adenosine triphosphatase. Science 250, 651–657.

Kravets, A.N., Zakharova, M.V., Solonin, A.S., Kuz'min, N.P., Taniashin, V.I., Glatman, L.I., Moroz, A.F., Baev, A.A., 1990. Cloning and regulation of gene expression of *Eco*RV restriction-modification system. Mol. Biol. 24, 438–447.

LaVallie, E.R., McCoy, J.M., 1995. Gene fusion expression systems in *Escherichia coli*. Curr. Opin. Biotechnol. 6, 501–506.

Lin, C.R., Kapiloff, M.S., Durgerian, S., Tatemoto, K., Russo, A.F., Hanson, P., Schulman, H., Rosenfield, M.G., 1987. Molecular cloning of a brain-specific calcium/calmodulin-dependent protein kinase. Proc. Natl. Acad. Sci. USA 84, 5962–5966.

Nilsson, B., Abrahmsen, L., 1990. Fusions to Staphylococcal protein A. Methods Enzymol. 185, 144–161.

Page, A.P., Landry, D., Wilson, G.G., Carlow, C.K.S., 1995. Molecular characterization of a cyclosporin A-insensitive cyclophilin from the parasitic nematode *Brugia malayi*. Biochemistry 34, 11545–11550.

Perler, F.B., Comb, D.G., Jack, W.E., Moran, L.S., Qiang, B., Kucera, R.B., Benner, J., Slatko, B.E., Nwankwo, D.O., Hempstead, S.K., Carlow, C.K.S., Jannasch, H., 1992. Intervening sequences in an Archaea DNA polymerase gene. Proc. Natl. Acad. Sci. USA 89, 5577–5581.

Perler, F.B., Davis, E.O., Dean, G.E., Gimble, F.S., Jack, W.E., Neff, N., Noren, C.J., Thorner, J., Belfort, M., 1994. Protein splicing elements: inteins and exteins a definition of terms and recommended nomenclature. Nucleic Acids Res. 22, 1125–1127.

Roberts, G.A.F. and Taylor, K. (1988) The preparation and characterization of chitin beads for the use in chromatography. In: Skyjåk-Bræk, G., Anthonsen, T. and Sanford, P. (Eds.), Chitin and Chitosan Sources, Chemistry, Biochemistry, Physical Properties, and Applications. Elsevier Applied Science, London, pp. 577–583.

Shao, Y., Xu, M.-Q., Paulus, H., 1995. Protein splicing: characterization of the aminosuccinimide residue at the carboxyl terminus of the excised intervening sequence. Biochemistry 34, 10844–10850.

Shao, Y., Xu, M.-Q., Paulus, H., 1996. Protein splicing: evidence for an N-O acyl rearrangement as the initial step in the splicing process. Biochemistry 35, 3810–3815.

Shumacher, T.N.M. and Tsomides, T.J. (1996) In vitro radiolabeling of peptides and proteins. In: Coligan, J.E., Dunn, B.M., Ploegh, H.L., Speicher, D.W. and Wingfield, P.T. (Eds.), Current Protocols in Protein Science, Vol. 1. John Wiley and Sons, New York, pp. 3.3.1–3.3.19.

Smith, D.B., Johnson, K.S., 1988. Single-step purification of polypeptides expressed in *Escherichia coli* as fusions with glutathione *S*-transferase. Gene 67, 31–40.

Stofko-Hahn, R.E., Carr, D.W., Scott, J.D., 1992. A single step purification for recombinant proteins. Characterization of a microtubule associated protein (MAP 2) fragment which associates with the type II cAMP-dependent protein kinase. FEBS Lett. 302, 274–278.

Studier, F.W., Rosenberg, A.H., Dunn, J.J., Dubendorff, J.W., 1990. Use of T7 RNA polymerase to direct expression of cloned genes. Methods Enzymol. 185, 60–89.

Van Dyke, M.W., Sirito, M., Sawadogo, M., 1992. Single-step purification of bacterially expressed polypeptides containing an oligo-histidine domain. Gene 111, 99–104.

Watanabe, T., Ito, Y., Yamada, T., Hashimoto, M., Sekine, S., Tanaka, H., 1994. The role of the C-terminal domain and type III domains of chitinase A1 from *Bacillus circulans* WL-12 in chitin degradation. J. Bacteriol. 176, 4465–4472.

Wilkie, S.D., Landry, D., 1988. Determination of certain agarose immobilized ligands by the use of perchloric acid. BioChromatography 3, 205–214.

Xu, M.-Q., Perler, F.B., 1996. The mechanism of protein splicing and its modulation by mutation. EMBO J. 15, 5146–5153.

Xu, M.-Q., Southworth, M.W., Mersha, F.B., Hornstra, L.J., Perler, F.B., 1993. In vitro protein splicing of purified precursor and the identification of a branched intermediate. Cell 75, 1371–1377.

Xu, M.-Q., Comb, D.G., Paulus, H., Noren, C.J., Shao, Y., Perler, F.B., 1994. Protein splicing: an analysis of the branched intermediate and its resolution by succinimide formation. EMBO J. 13, 5517–5522.